MAT224 - LEC5101 - Lecture 4
Linear transformations and matrices, kernel and image

Dylan Butson

University of Toronto

January 28, 2020

Overview of today

- Linear algebra is about solving systems of linear equations:
- Fix $a_{ij} \in \mathbb{R}$ for $i = 1, ..., m$ and $j = 1, ..., n$
  Question: For each $\mathbf{b} = (b_1, ..., b_m)$, does there exist $\mathbf{x} = (x_1, ..., x_n)$ such that
  $$a_{i1}x_1 + ... + a_{in}x_n = b_i \qquad \text{for each } i = 1, ..., m?$$
- If no, then for which $\mathbf{b}$ does there exist such an $\mathbf{x}$?
- If yes, how many 'different' $\mathbf{x}$ are there for fixed $\mathbf{b}$?
- In 223, you learned algorithms to answer these questions.
- You also learned some geometric intuition for how these work, in terms of vectors in $\mathbb{R}^n$ and linear transformations $\mathbb{R}^n$ to $\mathbb{R}^m$.
- In 224, we're developing language to formalize this geometric intution into precise reasoning, which allows us to answer these questions without doing as much computation.
- We also consider 'general' vector spaces V, not just $\mathbb{R}^n$.

Discussion: Coordinates with respect to a basis

Let V be a vector space and $\beta = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ be a basis for V.

Recall that we proved:

Proposition: For every vector $\mathbf{v} \in V$, there **exists** a **unique**

$$x_1, ..., x_n \in \mathbb{R} \quad \text{such that} \quad \mathbf{v} = x_1\mathbf{v}_1 + ... + x_n\mathbf{v}_n$$

Notation: We call $(x_1, ..., x_n)$ the *coordinate vector* of $\mathbf{v}$ with respect to α, and write $[\mathbf{v}]^\alpha = [x_i]$ as a column vector.

Discussion: Let $V = \mathbb{R}^2$ and $\alpha = \{(1, -1), (1, 1)\}$. Determine

(1) $[\mathbf{e}_1]_\alpha$

(2) $[\mathbf{e}_2]_\alpha$

(3) $[\mathbf{e}_1 + \mathbf{e}_2]_\alpha$

where $\mathbf{e}_1 = (1, 0), \mathbf{e}_2 = (0, 1)$ are the standard basis vectors.

Let $V = P_2(\mathbb{R})$ and $\alpha = \{1, 1 + x, 1 + x + x^2\}$. Determine

(a) $[1]_\alpha$

(b) $[x]_\alpha$

(c) $[x^2]_\alpha$

(d) $[a + bx + cx^2]_\alpha$ where a, b, c are arbitrary scalars.

(e) Show that $[a + bx + cx^2]_\alpha = a[1]_\alpha + b[x]_\alpha + c[x^2]_\alpha$.

Discussion: Linear Transformations

Throughout, let S, S' be sets and V, W be vector spaces.

Definition: A *function* $f : S \to S'$ is a rule that assigns to each $s \in S$ and element $f(s) \in S'$; we write $s \mapsto f(s)$.

The set S is called the *domain* of f, and the set S' the *target* of f.

Definition: A function $T : V \to W$ is called *linear* if

$$T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}) \quad \text{and} \quad T(c\mathbf{x}) = cT(\mathbf{x})$$

for each $\mathbf{x}, \mathbf{y} \in V$, $c \in \mathbb{R}$. Equivalently, $T(c\mathbf{x} + \mathbf{y}) = cT(\mathbf{x}) + T(\mathbf{y})$.

Example: $T : \mathbb{R} \to \mathbb{R}$ defined by $T(x) = x^n$ is linear iff $n = 1$.

Discussion: Which of the following are linear? (with proof)

(1) $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T(\mathbf{x}) = A\mathbf{x}$ for $A \in M_{m\times n}(\mathbb{R})$.

(2) $T : P_n(\mathbb{R}) \to P_n(\mathbb{R})$ defined by $T(p) = \frac{d}{dx}p$.

(3) $T : M_{m\times n}(\mathbb{R}) \to M_{m\times r}(\mathbb{R})$ by $T(A) = AB$ for $B \in M_{n\times r}(\mathbb{R})$.

(4) Let $S = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ be a basis for V. Show the following:
For each $\mathbf{y}_1, ..., \mathbf{y}_n \in W$ there exists a unique linear $T : V \to W$ such that $T(\mathbf{v}_j) = \mathbf{y}_j$ for each $j = 1, ..., n$.

Matrices and Linear Transformations

Let V, W be vector spaces and $\alpha = \{\mathbf{v}_1, ..., \mathbf{v}_n\}, \beta = \{\mathbf{w}_1, ..., \mathbf{w}_m\}$ be bases for them. Thus, $\dim V = n, \dim W = m$.

Let $T : V \to W$ be a linear transformation. We just proved that T is determined uniquely by the vectors $T(\mathbf{v}_j) \in W$ for $j = 1, ..., n$.

For each j, the vector $T(\mathbf{v}_j) \in W$ has a unique decomposition

$T(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + ... + a_{mj}\mathbf{w}_m$ for some $a_{ij} \in \mathbb{R}$ for $i = 1, ..., m$.

As a column vector, we have $$[T(\mathbf{v}_j)]^\beta = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \in \mathbb{R}^m$$

In summary, given α, β we can record the information of T by:

$$[T]_\alpha^\beta = \big[T(\mathbf{v}_1) | \quad \cdots \quad | T(\mathbf{v}_n) \big]^\beta = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in M_{m \times n}(\mathbb{R})$$

The j^{th} column of $[T]_\alpha^\beta$ describes $T(\mathbf{v}_j)$, the image under T of the j^{th} vector $\mathbf{v}_j$ in the basis α, in terms of coordinates defined by β.

Discussion: Matrices and Linear Transformations

Let $V = \mathbb{R}^2$, $W = \mathbb{R}^3$, $\alpha = \{\mathbf{e}_1, \mathbf{e}_2\}$, $\beta = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ the standard bases, and $\alpha' = \{\mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2\}, \beta' = \{\mathbf{f}_1 + \mathbf{f}_2, \mathbf{f}_1 - \mathbf{f}_2, \mathbf{f}_2 + \mathbf{f}_3\}$.
Define a linear map $T : V \to W$ by

$T(\mathbf{e}_1) = \mathbf{f}_1 + \mathbf{f}_2 \quad T(\mathbf{e}_2) = \mathbf{f}_2 + \mathbf{f}_3$ and calculate

(1) the matrix $[T]_\alpha^\beta$
(2) the matrix $[T]_{\alpha'}^\beta$
(3) the matrix $[T]_\alpha^{\beta'}$

Let $V = P_2(\mathbb{R})$, and $\alpha = \{1, x, x^2\}, \beta = \{1 + x, 1 - x, x^2\}$.
Define $T : V \to V$ by $T(p) = \frac{d}{dx} p$ and calculate

(4) the matrix $[T]_\alpha^\alpha$
(5) the matrix $[T]_\alpha^\beta$
(6) the matrix $[T]_\beta^\alpha$

Bonus: Let V, W vector spaces, α, β bases, and $T : V \to W$.
Prove that $\quad [T(\mathbf{x})]^\beta = [T]_\alpha^\beta \cdot [\mathbf{x}]^\alpha \quad$ for each $\mathbf{x} \in V$.
This is just showing that '**matrix multiplication works**'.

Discussion: Injective and Surjective Functions

Let $S, \tilde{S}$ be sets, $R \subset S, \tilde{R} \subset \tilde{S}$ be subsets and $f : S \to \tilde{S}$.

Definition:

- The *image of R under f* is $f(R) = \{f(s) | s \in R\} \subset \tilde{S}$.
- The *preimage of $\tilde{R}$ under f*, $f^{-1}(\tilde{R}) = \{s \in S | f(s) \in \tilde{R}\} \subset S$

Warning: The preimage is always defined even if f is not invertible.

Definition: f is *injective* if knowing $f(s) = f(t)$ implies $s = t$.

f is *surjective* if for each $\tilde{s} \in \tilde{S}$ there exists $s \in S$ with $f(s) = \tilde{s}$.

f is *bijective* if it is injective and surjective.

Discussion: Prove the following:

(1) f is surjective if and only if $f(S) = \tilde{S}$, if and only if:
For each $\tilde{s} \in \tilde{S}$, $f^{-1}(\{\tilde{s}\})$ is non-empty, i.e. $f^{-1}(\{\tilde{s}\}) \neq \emptyset$.

(2) f is injective if and only if:
For each $\tilde{s} \in \tilde{S}$, $f^{-1}(\{\tilde{s}\})$ is either a single point $\{s\}$ or empty $\emptyset$.

(3) f is bijective if and only if:
For each $\tilde{s} \in \tilde{S}$, there **exists** a **unique** $s \in S$ such that $f(s) = \tilde{s}$.

Kernel and Image

Let V, W be vector spaces and $T : V \to W$ a linear map.

Definition: The *kernel* of T is the subset of V defined by

$$\ker(T) = T^{-1}(\{\mathbf{0}\}) = \{\mathbf{v} \in V | T(\mathbf{v}) = \mathbf{0}\}$$

Example: Define $T : \mathbb{R}^3 \to \mathbb{R}^2$ by $T(x_1, x_2, x_3) = (x_1, x_2 - x_3)$.

$$\ker(T) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_1 = 0,\ x_2 = x_3\}$$

Definition: The *image* of T is the subset of W defined by

$$\operatorname{im}(T) = T(V) = \{T(\mathbf{v}) \in W | \mathbf{v} \in V\}$$

Example: Define $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T(x_1, x_2) = (x_1, x_2, x_1 - x_2)$.

$$\operatorname{im}(T) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_3 = x_1 - x_2\}$$

Discussion: Kernel and Image

Let V, W be vector spaces with bases $\alpha = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ and $\beta = \{\mathbf{w}_1, ..., \mathbf{w}_m\}$, and $T : V \to W$ a linear map.

Prove the following:

(1) $\ker(T)$ is a subspace of V.

(2) $\mathrm{im}(T)$ is a subspace of W.

(3) $\mathrm{im}(T) = \mathrm{Span}(\{T(\mathbf{v}_1), ..., T(\mathbf{v}_n)\})$

(4) $\ker(T) = \mathrm{null}([T]_\alpha^\beta)$ (use that 'matrix multiplication works')

(5) $\mathrm{im}(T) = \mathrm{col}([T]_\alpha^\beta)$ (use that 'matrix multiplication works')

(6) T is injective if and only if $\ker(T) = \{\mathbf{0}\}$

Bonus: Let $\mathbf{b} \in \mathrm{im}(T)$ so that $\mathbf{b} = T(\mathbf{x}_0)$ for $\mathbf{x}_0 \in V$. Then show

$$T^{-1}(\{\mathbf{b}\}) := \{\mathbf{x} \in V | T(\mathbf{x}) = \mathbf{b}\} = \{\mathbf{x}_0 + \mathbf{v} | \mathbf{v} \in \ker(T)\}$$

Conclude that there is a bijection between $T^{-1}(\{\mathbf{b}\})$ and $\ker(T)$.

What does it mean to solve linear equations?

Fix $a_{ij} \in \mathbb{R}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

Question: For each $\mathbf{b} = (b_1, \ldots, b_m)$, does there exist $\mathbf{x} = (x_1, \ldots, x_n)$ such that

$$a_{i1}x_1 + \ldots + a_{in}x_n = b_i \qquad \text{for each } i = 1, \ldots, m?$$

- If no, then for which $\mathbf{b}$ does there exist such an $\mathbf{x}$?
- If yes, how many 'different' $\mathbf{x}$ are there for fixed $\mathbf{b}$?

Answer: Let $A = [a_{ij}] \in M_{m \times n}(\mathbb{R})$, which defines $T : \mathbb{R}^n \to \mathbb{R}^m$. Then

$$a_{i1}x_1 + \ldots + a_{in}x_n = b_i \qquad \text{for each } i = 1, \ldots, m$$

if and only if

$$A\mathbf{x} = \mathbf{b}, \text{ or equivalently } T(\mathbf{x}) = \mathbf{b}.$$

Thus, we have the following answer:

- There exists $\mathbf{x}$ solving the equation if and only if $\mathbf{b} \in \operatorname{im}(T)$.
- For each fixed $\mathbf{b}$, the set of solutions is $T^{-1}(\{\mathbf{b}\})$, which we showed is in bijection with $\ker(T)$.

Towards The Dimension Theorem

We have reduced the question of **existence** and **uniqueness** of solutions to linear equations to understanding the **image** and **kernel** of a linear map $T : V \to W$.

What can we say about $\mathrm{im}(T)$ and $\ker(T)$ in general? Let's look at some examples:

(1) $T : \mathbb{R}^2 \to \mathbb{R}^3 \quad (x_1, x_2) \mapsto (x_1, x_2, 0)$
- $\dim \ker(T) = 0$, $\dim \mathrm{im}(T) = 2$.

(2) $T : \mathbb{R}^2 \to \mathbb{R}^3 \quad (x_1, x_2) \mapsto (x_1, 0, 0)$
- $\dim \ker(T) = 1$, $\dim \mathrm{im}(T) = 1$.

(3) $T : \mathbb{R}^2 \to \mathbb{R}^3 \quad (x_1, x_2) \mapsto (0, 0, 0)$
- $\dim \ker(T) = 2$, $\dim \mathrm{im}(T) = 0$.

(4) $T : \mathbb{R}^n \to \mathbb{R}^m \quad (x_1, ..., x_n) \mapsto (x_1, ..., x_{n-k}, 0, ..., 0)$
- $\dim \ker(T) = k$, $\dim \mathrm{im}(T) = n - k$.

Claim: Every linear $T : V \to W$ looks like this wrt some bases.

Corollary: Let $T : V \to W$ linear, with $\dim V = n$. Then

$$\dim V = n = k + (n - k) = \dim \ker(T) + \dim \mathrm{im}(T)$$

Discussion: The Dimension Theorem

Theorem: (The Dimension Theorem) Let $T : V \to W$ be a linear map, with V finite dimensional. Then

$$\dim V = \dim \ker(T) + \dim \operatorname{im}(T)$$

Let's prove the dimension theorem using the following steps:

To fix notation, let's say $\dim V = n$.

(1) Since $\ker(T) \subset V$, we know $k := \dim \ker(T) \leq \dim V = n$.

(2) Choose a basis $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ for $\ker(T)$, and extend to a basis $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$ for V.

(3) Show that $\{T(\mathbf{v}_{k+1}), ..., T(\mathbf{v}_n)\}$ define a basis for $\operatorname{im}(T)$. (If it were linearly dependent, find a 'new' element of $\ker(T)$)

(4) Conclude that $\dim \operatorname{im}(T) = n - k$.

(5) Use that $n = k + (n - k)$ to prove the theorem.

Suppose $\dim W = m$ is finite, and extend $\{T(\mathbf{v}_{k+1}), ..., T(\mathbf{v}_n)\}$ to a basis for W.

What is $[T]$ with respect to these bases?